Huberian approach for reduced order ARMA modeling of neurodegenerative disorder signal

Christophe Corbier, Mohamed El Badaoui, Hector Manuel Romero Ugalde

Huberian Approach for Reduced Order ARMA Modeling of Neurodegenerative Disorder Signal

Christophe Corbier[a,*], Mohamed El Badaoui[a], Hector Manuel Romero Ugalde[b]

[a] Universite de Lyon, F-42023 Saint Etienne, France
Universite de Saint Etienne, Jean Monnet, F-42000 Saint-Etienne, France
LASPI, F-42334 IUT de Roanne, France
[b] Universite de Rennes 1, 35042, Rennes, France
LTSI

Abstract

The purpose of this paper is to address the question of the existence of auto regressive moving average (ARMA) models with reduced order for neurodegenerative disorder signals by using Huberian approach. Since gait rhythm dynamics between Parkinson's disease (PD) or Huntington's disease (HD) and healthy control (CO) differ, and since the stride interval presents great variability, we propose a different ARMA modeling approach based on a Huberian function to assess parameters. Huberian function as a mixture of L_2 and L_1 norms, tuned with a threshold γ from a new curve, is chosen to deal with stride signal disorders. The choice of γ is crucial to ensure a good treatment of NO and allows to reduce the model order. The disorders induce disturbances in the classical estimation methods and increase of the number of parameters of the ARMA model. Here, the use of the Huberian function reduces the number of parameters of the estimated models leading to a *disease transfer function* with low order for PD and HD. Mathematical approach is discussed and experimental results based on a database containing 16 CO, 15 PD, and 19 HD are presented.

Keywords: Reduced order ARMA model, Gait signal, Huberian function, Tuning function, L_1 contribution, Neurodegenerative disease

1. Introduction

This paper introduces a new parametric approach for the estimation problem of the reduced order auto regressive moving average (ROARMA) model of human gait rhythm signal [13]. ARMA system identification is a well-defined problem in several science and engineering areas such as speech signal processing, adaptive filtering, radar Doppler processing or biomechanics. There exists different methods to deal with the ARMA estimation problem. Based on the fractional signal processing approach, Chaudhary et al [11] proposes a fractional least mean square (LMS) algorithm for parameter estimation of Hammerstein nonlinear ARMA system with exogenous noise.

*Corresponding author. Tel.: +(33)477448157
E-mail address: christophe.corbier@univ-st-etienne.fr

This algorithm has still been used in other studies [2] [41] [10]. Another approach uses a two-stage fractional LMS identification algorithm for parameter estimation of controlled ARMA (CARMA) systems [33]. The main idea is to use fractional LMS identification (FLMSI) and two-stage FLMSI (TS-FLMSI) algorithms for CARMA models which are decomposed into a system and noise models. Based on robust estimation, Chakhchoukh [9] introduces a new robust method to estimate the parameters of a Gaussian ARMA model contaminated with outliers [18]. This method makes use of a median and is termed ratio-of-medians estimator (RME). Among the problems of ARMA identification, the model order estimation is crucial. Al-Qawasmi et al [4] propose a new technique for model to estimate order in a general ARMA process based on a rounding approach. Most of the time, these estimation procedures are performed by the implicit assumption that the processes are Gaussian [34]. However, most real world signals are non-Gaussian and different methods such as higher order statistics are used [3] [40]. Moreover, these methods are based on the assumption that the signal does not contain outliers or a low density of outliers less than 1%. A reference paper in a robust estimation framework uses Huberian function for ARMA models [30]. This work shows that the Huberian-estimates are closely related to those based on a robust filter, but they have two important advantages: they are consistent and the asymptotic theory is tractable. However, in this analysis, the residuals are computed so the effect of one outlier is limited to the period where it occurs. Moreover, experimental results only focus on the Monte Carlo simulations, not real measurements. A recent paper [45] developed a systematic procedure of statistical inference for the ARMA model with unspecified and heavy-tailed heteroscedastic noises. The authors compare some estimators such that LSE, Huberian function and generalized Huberian function with outliers in a simulated ARMA process. In our framework, the measurements are real and contain natural outliers (NO) due to the neurodegenerative disorders of each disease.

Neurodegenerative disorders have a direct consequence on the human behavior by introducing NO in biomechanic time-signals. These points are crucial in the study of neurodegenerative diseases and provide information of the degree of disorder. Here, the Parkinson's disease (PD) and Huntington disease (HD) are studied through the *stride time-signal* (STS) of human gait rhythm, corresponding to the time from initial contact of when one foot to the subsequent contact of the same foot [21]. Walking is one of the most fundamental and important activities of human that is strongly related to human health [39]. This is a complex process which we have only recently begun to understand through the study of the interval data in a complete gait cycle [35] [36]. Gait rhythm can also be described in terms of swing and stance intervals corresponding to the time of one foot is in the air and the time of bilateral foot contact, respectively (Fig.A.1). Human locomotion is regulated by the central nervous system (CNS). In the CNS of the human body, motor neurons are the nerve cells that process sensory information and control voluntary muscle movement [37]. Serving as a pivotal part of the human motor system, the basal ganglia process motor impulses originating from the cerebral cortex and the brain stem, and also sends sensory information through the projecting loops in the CNS [42]. Basal ganglia dysfunction affects motor function and may lead to balance impairment or altered gait rhythm. PD is a chronic and progressive hypokinetic disorder of the CNS induced by basal ganglia dysfunction. HD is a progressive neurodegenerative disorder with autosomal dominant inheritance. Analysis of gait parameters is very useful for a better understanding of the mechanisms of movement disorders, in particular for neurodegenerative diseases.

Different approaches exist to analyze gait rhythm time-signals, such as the kinematic aspect [29] [24], Gaussian approach [43] [23], Huberian framework [13], and cyclostationary analysis [28] [44]. Wu and Krishnan [43] developed a framework through Gaussian statistical analysis applied to PD, amyotrophic lateral sclerosis, and gait maturation in children. The main drawback of studies based on the Gaussian framework is the not well treatment of the NO in the time-signal. Indeed, during the 5-min walking period, every time the subjects reached the end of the hallway, they had to turn around, and finally they continued walking. The time-signal stride recorded during these walking turns should be treated as NO. The authors replaced these points by the median value of the stride interval time series, using the *three-sigma rule*, in order to avoid disturbance of the statistical moments. Unfortunately, these authors neglected relevant information about the time-signal dynamics, since these NO give capital information during the short phase of the walking turn. These subjects present difficulties to turn and it seems fundamental to consider these points. Therefore, Gaussian-based estimation cannot be applied.

Here we propose a reduced order ARMA modeling approach based on a Huberian function to assess parameters and experimental results are performed with STS real measurements of CO, PD and HD. Huberian function is a mixture of L_2 and L_1 norms with a threshold γ. The choice of γ is crucial to ensure a good treatment of NO and allows to reduce the model order. A large section in this paper discusses on the choice of γ using a new curve. A relevant choice of γ in a new interval range ensures both convergence and consistency of the robust estimator. Convergence is shown and a new method to assess the variance/covariance matrix of the estimator is proposed.

This paper is organized as follows: Section 2 gives the Huberian mathematical context of the ARMA estimator. Experimental results based on a database containing 16 CO, 15 PD, and 19 HD are shown in Section 3. Conclusions and perspectives are drawn in Section 4.

2. Huberian mathematical framework

This section presents the Huberian framework mathematical basis. The choice of the threshold in Huber's function is presented and discussed. Asymptotic convergence in law of the robust estimator is shown, considering the stochastic differentiability approach [31] and the m-dependence context. A new method to assess the variance/covariance matrix of the estimator is proposed.

2.1. Huberian function and estimation criterion

Let $(S, \mathbf{S}, P)$ be a probability space and $\{X_k\}_{k=1}^{N}$ a sequence of i.i.d.r.v's with values in S. Let Θ be a Borel subset in $\mathbb{R}^d$ and Γ a compact subset of $\mathbb{R}$. Let $\rho_\gamma^H : S \times \Theta \times \Gamma \to \mathbb{R}$ be a symmetric function such that $\rho_\gamma^H(\bullet(\theta,\gamma))$ is measurable for each $\theta \in \Theta$ and $\gamma \in \Gamma$. The estimator $\hat{\theta}_N^H$ is defined by a minimum of the form

$$N^{-1}\sum_{k=1}^{N}\rho_\gamma^H\left(X_k(\hat{\theta}_N^H,\hat{\gamma})\right) = \inf_{\theta\in\Theta,\gamma\in\Gamma} N^{-1}\sum_{k=1}^{N}\rho_\gamma^H\left(X_k(\theta,\gamma)\right) \tag{1}$$

with

$$\rho_\gamma^H(X) = \begin{cases} \frac{X^2}{2} & \text{for } |X| \le \gamma \\ \gamma|X| - \frac{\gamma^2}{2} & \text{for } |X| > \gamma \end{cases} \tag{2}$$

where γ is a threshold to be determined to improve efficiency, convergence, and stability of $\hat{\theta}_N^H$ [22] [12]. Let us introduce two index sets in $\theta \in \mathbb{R}^d$ defined by $\nu_2(\theta,\gamma) = \{k : |\varepsilon_k(\theta,\gamma)| \le \gamma\}$ and $\nu_1(\theta,\gamma) = \{k : |\varepsilon_k(\theta,\gamma)| > \gamma\}$ such that $card\left[\nu_2(\theta,\gamma)\right] + card\left[\nu_1(\theta,\gamma)\right] = N \ \forall\theta \in \mathcal{D}_M, \ \gamma \in \mathcal{D}_\gamma$, where $\mathcal{D}_M$ and $\mathcal{D}_\gamma$ are compact subsets and M a model structure. Let $M(\theta)$ be a particular model corresponding to the parameter vector value θ. Let us define $\tilde{\theta} = [\theta \ \ \gamma]$. Let $W_N(\theta,\gamma)$ be the estimation criterion of the parameter vector θ for a threshold $\gamma > 0$. We denote $s_k(\theta,\gamma)$, $k = 1,...,N$ the sign function such that $s_k(\theta,\gamma) = 1$ for $\varepsilon_k(\theta,\gamma) > \gamma$, $s_k(\theta,\gamma) = -1$ for $\varepsilon_k(\theta,\gamma) < -\gamma$ and $s_k(\theta,\gamma) = 0$ for $|\varepsilon_k(\theta,\gamma)| < \gamma$. Let $\varepsilon_k(\theta,\gamma) = y_k - \hat{y}_{k|k-1}(\theta,\gamma) = y_k - \varphi_k^T(\theta,\gamma)\theta$ be the prediction error where y_k is the process output, $\hat{y}_{k|k-1}(\theta,\gamma)$ the prediction model and $\varphi_k(\theta,\gamma) \in \mathbb{R}^d$ the regressor vector. This criterion contains a L_2 part to treat small prediction errors and a L_1 part to deal with NO. Consider a batch of data from the system $\tilde{Z}^N = [y_1...y_N]$. Roughly speaking, we have to determine a mapping from the data $\tilde{Z}^N$ to the set $\mathcal{D}_M \times \mathcal{D}_\gamma$

$$\tilde{Z}^N \longrightarrow \hat{\tilde{\theta}}_N^H = \left[\hat{\theta}_N^H \ \ \hat{\gamma}\right] \in \mathcal{D}_M \times \mathcal{D}_\gamma \tag{3}$$

The robust estimation criterion can be written as

$$W_N(\theta,\gamma) = \frac{1}{N}\sum_{k\in\nu_2(\theta,\gamma)}\frac{\varepsilon_k^2(\theta,\gamma)}{2} + \frac{\gamma}{N}\sum_{k\in\nu_1(\theta,\gamma)}\left(|\varepsilon_k(\theta,\gamma)| - \frac{\gamma s_k^2(\theta,\gamma)}{2}\right) \tag{4}$$

Let us denote $\|X\|^2 = \sum_i x_i^2$ and $|X| = \sum_i |x_i|$ where $X = [x_1...x_N]^T$. We define the following rule: $x_{\nu_i,k} = x_k$ for all $k \in \nu_i(\theta,\gamma)$ and $x_{\nu_i,k} = 0$ otherwise. We define the sparse matrix in $\mathbb{R}^{N\times d}$ over $\nu_i(\theta,\gamma)$ $(i = 1,2)$ respectively given by

$$\Phi_{\nu_i}(\theta,\gamma) = \begin{bmatrix} \varphi_{\nu_i,1}^T(\theta,\gamma) \\ ... \\ \varphi_{\nu_i,N}^T(\theta,\gamma) \end{bmatrix}, \quad \varphi_{\nu_i,k}(\theta,\gamma) = \begin{cases} \varphi_k(\theta,\gamma) & \text{for } \ k \in \nu_i(\theta,\gamma) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

On the other hand, we define $Y_{\nu_i} = [y_{\nu_i,1}...y_{\nu_i,N}]^T$ the process output vector and $S_{\nu_1} = [s_{\nu_1,1}...s_{\nu_1,N}]^T$ the sign vector. The estimation criterion to be minimized is then given by

$$W_N(\theta,\gamma) = \frac{1}{2N}\left\|Y_{\nu_2} - \Phi_{\nu_2}(\theta,\gamma)\theta\right\|^2 + \frac{1}{N}\left[\gamma\left|Y_{\nu_1} - \Phi_{\nu_1}(\theta,\gamma)\theta\right| - \frac{\gamma^2}{2}\left\|S_{\nu_1}\right\|^2\right] \tag{6}$$

This minimization algorithm is applied to yield a minimum corresponding to a given robust estimator for an appropriated choice of the threshold γ. In the sequel, we show this choice from two joint approaches. The first one comes from the maximum of the bias by defining a new function with properties to reduce the effect of NO in prediction errors. A new curve is presented and locates a new investigation interval of γ. From this, the second approach is to seek a local or global minimum of the robust estimation criterion with respect to θ and γ.

2.2. ARMA model in Huber's framework

The process output data are denoted as δt_k, $k = 1...N$ corresponding to the STS of human gait rhythm. Figure A.1 shows an example of the left gait signal from heel toe force sensors underneath the left foot where appear the different phases. Now assuming that δt_k is generated according to

$$\delta t_k = H_0(q)\, e_k \tag{7}$$

where $H_0(q)$ is the noise filter and e_k, $k = 1...N$ a random variables sequence with zero mean and variances λ. The ARMA model set is parametrized by a d-dimensional real-valued parameter vector θ, i.e.,

$$\delta t_k = H(q,\theta)\, e_k = \frac{C(q,\theta)}{\mathcal{A}(q,\theta)} e_k \tag{8}$$

with $\mathcal{A}(q,\theta) = 1 + \sum_{i=1}^{n_A} a_i q^{-i}$, $C(q,\theta) = 1 + \sum_{i=1}^{n_C} c_i q^{-i}$ and $\theta = [a_1...a_{n_A} c_1...c_{n_C}]^T$. Moreover, q^{-1} is the lag operator such that $q^{-l}\delta t_k = \delta t_{k-l}$, $l \in \mathbb{N}$.

In Huber's framework, the prediction errors depends on θ and γ. We write $\varepsilon_k(\theta,\gamma) = \delta t_k - \hat{\delta t}_k(\theta,\gamma)$ where $\hat{\delta t}_k(\theta,\gamma) = \varphi_k^T(\theta,\gamma)\,\theta$ is the prediction model. The regressor is $\varphi_k(\theta,\gamma) = [-\delta t_{k-1} ... -\delta t_{k-n_A} \;\; \varepsilon_{k-1}(\theta,\gamma) ... \varepsilon_{k-n_C}(\theta,\gamma)]^T$ and $\psi_k(\theta,\gamma)$ is the gradient with respect to θ of $\hat{\delta t}_k(\theta,\gamma)$ given by $\psi_k(\theta,\gamma) = \frac{1}{C(q,\theta)}\varphi_k(\theta,\gamma)$, meaning that $\psi_k(\theta,\gamma)$ is obtained by filtering the vector $\varphi_k(\theta,\gamma)$ through a stable linear filter.

2.3. Choice of γ

2.3.1. Location of γ

In the prediction error procedure, there appears an inner feedback loop to compute the pseudolinear prediction model $\hat{y}_{k|k-1}(\theta,\gamma)$. The estimated residuals are treated by a parametric adaptive algorithm which includes $W_N(\theta,\gamma)$

to be minimized. The presence of NO in the process output y_k induces large values in $\varepsilon_k(\theta,\gamma)$. A convenient choice of γ improves the robustness by reducing the effects of these large deviations. In the literature, γ is chosen in the interval range $[1,2]$ for linear models. However, this choice does not ensure convergence, consistency nor stability of $\hat{\theta}_N^H$. Accordingly, the probability density function (pdf) of $\varepsilon_k(\theta,\gamma)$ is strongly disturbed and presents heavy tails. It is shown that *H*uber's estimators are not always robust and efficient when $\gamma \in [1,2]$. In a recent paper [14] on piezoelectric-systems, the use of small values of γ in $[0.01, 0.5]$ led to derive relevant output error models. In this work, even though the prediction errors were disturbed by numerous NO, the choice of the small values of γ around 0.05 allowed to obtain interesting results in the frequency interval range for the vibration drilling control. In the sequel, we introduce a new curve ensuring a reduction of the bias and we show the choice of γ in low values. In [12] (chapter 6, p.130), we studied the quality of the robustness through influence function [19] of the robust estimator. We showed that the upper bound of the bias is proportional to the high NO, denoted $\mathcal{L}^p$ and a new function named *tuning function*, denoted $f^{\omega}(\gamma)$. Figure A.2 shows this curve. It appears the *classical interval*, denoted C_γ where $\gamma \in [1, 1.5]$ and a new interval, named *extended interval*, denoted E_γ where $\gamma \in [0.001, 0.2]$. We showed that

$$\sup_{F_N \in \mathcal{P}_{\Phi_N}(\omega)} \left|\hat{\theta}_N^H - \theta^*\right| = b_N^{\omega}(k) \leq \hat{\kappa}^N f^{\omega}(\gamma)\,|\mathcal{L}^p| \tag{9}$$

where κ^N is independent of γ, θ^* is the true parameter, $\mathcal{P}_{\Phi_N}(\omega)$ is the corrupted distribution model and F_N the contaminated Gaussian. An approximation can be written as $f^{\omega}(\gamma) \approx 0.034\gamma^5 - 0.316\gamma^4 + 1.113\gamma^3 - 1.773\gamma^2 + 1.088\gamma - 0.002$. From a linearization of $f^{\omega}(\gamma)$ in C_γ and E_γ, in absolute value, the slope in E_γ is six times as important as that of the slope in C_γ. Accordingly, the sensitivity to reduce the influence of high NO in E_γ is six times as important. Therefore, this new curve allows to locate a new investigation interval of γ in low values in order to get low values of $f^{\omega}(\gamma)$ to decrease the effects of NO.

2.3.2. *Convergence domain of γ*

Consider the differential of $W_N(\theta,\gamma)$ with respect to θ and γ given by

$$dW_N(\theta,\gamma) = \partial_\theta W_N(\theta,\gamma)d\theta + \partial_\gamma W_N(\theta,\gamma)d\gamma \tag{10}$$

where ∂_X is the derivative with respect to X. In detail

$$\partial_\theta W_N(\theta,\gamma) = \frac{-1}{N}\sum_{k\in\nu_2(\theta,\gamma)} \psi_k(\theta,\gamma)\varepsilon_k(\theta,\gamma) - \frac{\gamma}{N}\sum_{k\in\nu_1(\theta,\gamma)} \psi_k(\theta,\gamma)s_k(\theta,\gamma) \tag{11}$$

with $\psi_k(\theta,\gamma) = -\partial_\theta \varepsilon_k(\theta,\gamma)$ and

$$\partial_\gamma W_N(\theta,\gamma) = \frac{1}{N}\sum_{k\in\nu_2(\theta,\gamma)} \phi_k(\theta,\gamma)\varepsilon_k(\theta,\gamma) + \frac{1}{N}\sum_{k\in\nu_1(\theta,\gamma)} \left(|\varepsilon_k(\theta,\gamma)| - \gamma s_k^2(\theta,\gamma) + \gamma\phi_k(\theta,\gamma)s_k(\theta,\gamma) - \frac{\gamma^2}{2}\phi_k^*(\theta,\gamma)s_k(\theta,\gamma)\right) \tag{12}$$

with $\phi_k(\theta,\gamma) = \partial_\gamma \varepsilon_k(\theta,\gamma)$ and $\phi_k^*(\theta,\gamma) = \partial_\gamma s_k(\theta,\gamma)$. Let us define $\tilde{\Psi}(\theta,\gamma) = \frac{dW_N(\theta,\gamma)}{d\theta} = \left[\Psi(\theta,\gamma) \;\; \partial_\gamma W_N(\theta,\gamma)\right]^T$, where $\tilde{\Psi}(\theta,\gamma) \in \mathbb{R}^{d+1}$ and $\Psi(\theta,\gamma) = \partial_\theta W_N(\theta,\gamma)$ named Ψ-function.

We seek an optimal value of γ such that $W_N(\theta,\gamma)$ presents a global minimum with probability one (w.p.1) as N tends to infinity, denoted $\overline{W}(\theta,\gamma) = \lim_{N\to\infty} EW_N(\theta,\gamma)$. This involves that the solution of $\tilde{\Psi}\left(\hat{\theta}_N^H, \hat{\gamma}\right) = 0$ is unique. However, it may happen that $\overline{W}(\theta,\gamma)$ does not have a unique global minimum, then we define two compact subsets $\mathcal{D}_c^\theta$ and $\mathcal{D}_c^\gamma$ such that $\hat{\theta}_N^H \to \mathcal{D}_c^\theta$ w.p.1 as $N\to\infty$ and $\hat{\gamma} \to \mathcal{D}_c^\gamma$. We then have

$$\hat{\tilde{\theta}}_N^H = \left[\hat{\theta}_N^H \;\; \hat{\gamma}\right] \to \mathcal{D}_c^\theta \times \mathcal{D}_c^\gamma \;\text{ w.p.1 as }\; N\to\infty \tag{13}$$

If we denote $\mathcal{D}_c^{\theta\gamma} = \mathcal{D}_c^\theta \times \mathcal{D}_c^\gamma$ then

$$\mathcal{D}_c^{\theta\gamma} = \operatorname*{argmin}_{\theta\in\mathcal{D}_M \; \gamma\in\mathcal{D}_\gamma} \overline{W}(\theta,\gamma) = \left\{\theta\in\mathcal{D}_M, \gamma\in\mathcal{D}_\gamma \;\; \overline{W}(\theta,\gamma) = \min_{\theta'\in\mathcal{D}_M \; \gamma'\in\mathcal{D}_\gamma} \overline{W}(\theta',\gamma')\right\} \tag{14}$$

theorem 1. *Consider a uniformly stable, linear model structure M. Assume that the data set* $\tilde{Z}^\infty = \lim_{N\to\infty} \tilde{Z}^N$*, then*

$$\sup_{\theta\in\mathcal{D}_M \; \gamma\in\mathcal{D}_\gamma} \left|W_N(\theta,\gamma) - \overline{W}(\theta,\gamma)\right| \to 0 \Rightarrow \inf_{\tilde{\theta}^*\in\mathcal{D}_c^{\theta\gamma}} \left|\hat{\tilde{\theta}}_N^H - \tilde{\theta}^*\right| \to 0 \;\; \textit{w.p.1 as } \; N\to\infty, \;\; \tilde{\theta}^* = [\theta^* \;\; \gamma^*] \tag{15}$$

See proof in ([12], chap.4 p.69). In the case where the condition $\tilde{\Psi}\left(\hat{\theta}_N^H, \hat{\gamma}\right) = 0$ does not present a unique solution, there exists a convergence domain of $\hat{\gamma}$ involving a local minimum of $\hat{\theta}_N^H$ such that $\hat{\gamma} \to \gamma^*$ and $\hat{\theta}_N^H \to \theta^*$ w.p.1 as $N\to\infty$. Using theorem 1 and $\inf_{\tilde{\theta}^*\in\mathcal{D}_c^{\theta\gamma}} \left|\hat{\tilde{\theta}}_N^H - \tilde{\theta}^*\right| \to 0$ w.p.1 as $N\to\infty$, the consistency of the robust estimator is proved.

Main properties of the robust estimator related to the covariance matrix and asymptotic normality of $\sqrt{N}\left(\hat{\theta}_N^H - \theta^*\right)$ are given. In the sequel we assume that $\hat{\gamma}$ converges to γ^* satisfying the conditions of theorem 1. Hence we suppose that the set $\mathcal{D}_c^{\theta\gamma}$ consists only one point $\tilde{\theta}^* = [\theta^* \;\; \gamma^*]$. We shall work with the expression $W_N(\theta,\gamma^*)$, $\theta\in\mathcal{D}_M$ and the derivatives will be carried out with respect to θ and will be denoted $\partial_\theta W_N(\theta,\gamma^*)$ and $\partial^2_{\theta\theta} W_N(\theta,\gamma^*)$ for the first and second derivatives respectively.

2.4. ML robust estimator

The robust estimator $\hat{\theta}_N^H$ is a maximum likelihood estimator (MLE) satisfying $\rho_\gamma^H(X,\gamma) \sim -log f_H(X,\gamma)$ where $f_H(X,\gamma)$ is the pdf defined by

$$f_H(X,\gamma) = \begin{cases} f_{L_2}(X,\gamma) = C(\gamma)\, e^{\frac{-X^2}{2\phi^2}} & \text{for } |X| \le \gamma \\ f_{L_1}(X,\gamma) = C(\gamma)\, e^{\frac{-\gamma|X|}{\phi^2} + \frac{\gamma^2}{2\phi^2}} & \text{for } |X| > \gamma \end{cases} \tag{16}$$

$C(\gamma) = \frac{1}{2(K_1(\gamma)+K_2(\gamma))}$ with

$$\begin{cases} K_1(\gamma) = e^{\frac{\gamma^2}{2\phi^2}} \frac{\phi^2}{\gamma} \Gamma\left(1, \frac{\gamma^2}{\phi^2}\right) & \text{for } |X| > \gamma \\ K_2(\gamma) = \frac{\phi}{\sqrt{2}} \left[\Gamma\left(\frac{1}{2}\right) - \Gamma\left(\frac{1}{2}, \frac{\gamma^2}{2\phi^2}\right)\right] & \text{for } |X| \le \gamma \end{cases} \tag{17}$$

$\Gamma(a)$ and $\Gamma(a,X)$ are respectively the complete and incomplete Euler's gamma functions. The parameter ϕ is the standard deviation of f_H and we can verify that $\forall X \in \mathbb{R}$, $f_H(X,\gamma) \ge 0$ and $\int_{\mathbb{R}} f_H(X,\gamma)\, dX = 1$, which ensure that f_H is a pdf.

2.5. Asymptotic covariance matrix of $\hat{\theta}_N^H$ in ARMA model

Since $\hat{\theta}_N^H$ minimizes $W_N(\theta,\gamma^*)$ then $\partial_\theta W_N(\hat{\theta}_N^H,\gamma^*) = 0$. Expanding this expression into Taylor's series around θ^* gives

$$\hat{\theta}_N^H - \theta^* = -\left[\overline{\partial^2_{\theta\theta} W}(\theta^*,\gamma^*)\right]^{-1} \partial_\theta W_N(\theta^*,\gamma^*) \tag{18}$$

where $\partial_\theta W_N(\theta^*,\gamma^*)$ is given by (11) and $\overline{\partial^2_{\theta\theta} W}(\theta^*,\gamma^*) = \lim_{N\to\infty} E\partial^2_{\theta\theta} W_N(\hat{\theta}_N^H,\gamma^*)$ is the symmetric non-negative definite $d \times d$ limit Hessian matrix with

$$\partial^2_{\theta\theta} W_N(\theta,\gamma^*) = \frac{-1}{N} \sum_{k\in\nu_2(\theta,\gamma^*)} \left(\partial_\theta \psi_k^T(\theta,\gamma^*)\varepsilon_k(\theta,\gamma^*) - \psi_k(\theta,\gamma^*)\psi_k^T(\theta,\gamma^*)\right) - \frac{\gamma}{N} \sum_{k\in\nu_1(\theta,\gamma^*)} \partial_\theta \psi_k^T(\theta,\gamma^*) s_k(\theta,\gamma^*) \tag{19}$$

See proof in ([12], chap.4 p.63). From (18) and for N sufficiently large, the asymptotic covariance matrix of the robust estimator is given by

$$\text{cov}\left(\hat{\theta}_N^H\right) \sim \frac{\left[\overline{\partial^2_{\theta\theta} W}(\theta^*,\gamma^*)\right]^{-1} Q(\theta^*,\gamma^*) \left[\overline{\partial^2_{\theta\theta} W}(\theta^*,\gamma^*)\right]^{-1}}{N} = \frac{\mathcal{P}(\theta^*,\gamma^*)}{N} \tag{20}$$

where $Q(\theta^*,\gamma^*) = \lim_{N\to\infty} NE\partial_\theta W_N(\theta^*,\gamma^*)\partial_\theta W_N(\theta^*,\gamma^*)^T$ is named Q-matrix.

Remark

For the user, having processed N data points and determined $\hat{\theta}_N^H$ and γ^*, we may use

$$\text{cov}\left(\hat{\theta}_N^H\right) = \frac{\left[\partial^2_{\theta\theta} W_N(\hat{\theta}_N^H,\gamma^*)\right]^{-1} Q\left(\hat{\theta}_N^H,\gamma^*\right) \left[\partial^2_{\theta\theta} W_N(\hat{\theta}_N^H,\gamma^*)\right]^{-1}}{N} \tag{21}$$

as an estimate of $\frac{\mathcal{P}(\theta^*,\gamma^*)}{N}$.

ARMA models involve a pseudolinear prediction model in $\hat{\delta}t_k(\theta,\gamma)$. On the other hand $\psi_k(\theta,\gamma) = \frac{1}{C(q,\theta)}\varphi_k(\theta,\gamma)$ meaning that the matrix $\partial_\theta\psi_k^T(\theta,\gamma^*)$ in (19) is not equal to zero. The main drawback is the infinite sum of Taylor's expansion of $\psi_k(\theta,\gamma^*)$ and $\partial_\theta\psi_k^T(\theta,\gamma^*)$, increasing the computational cost of the estimated covariance matrix (21). Here, we show the main results of our method to limit Taylor's expansion with a large order. For more details see ([12], chap.5 p.74) . After straightforward calculations, we have

$$\psi_k\left(\hat{\theta}_N^H,\gamma^*\right) = \sum_{m=0}^{\infty} A_m^N \varphi_{k-m}(\hat{\theta}_N^H,\gamma^*),\ \ A_m^N \leq 1 \tag{22}$$

with $A_m^N \approx -2\sum_{k=1}^{\mathcal{F}(n_C/2)} \tilde{\mu}_k \rho_k^{m-1} cos\left(\Omega_k^m\right)$, where $\Omega_k^m = \tilde{\theta}_k + (m-1)\tilde{\varphi}_k$ if n_C is an even number and $\Omega_k^m = l\pi$, $l = \{m, m-1, 1, 0\}$ if n_C is an odd number. $\mathcal{F}(n)$ is the nearest integer less than or equal to n. The coefficients $\tilde{\mu}_k,\ \rho_k,\ \tilde{\theta}_k,\ \tilde{\varphi}_k$ are given by the n_C-poles $\{\pi_k\}_{k=1}^{n_C} = \rho_k e^{j\tilde{\varphi}_k}$, where $\rho_k < 1$ for $k = 1...n_C$ and $k-th$ residue $Res\left(\tilde{\Phi};\pi_k\right) = \tilde{\mu}_k e^{j\tilde{\theta}_k}$ of the transfer function

$$\tilde{\Phi}\left(e^{j\omega},\theta\right) = 1 - \frac{1}{C\left(e^{j\omega},\theta\right)} = \frac{c_1 e^{j\omega(n_C-1)} + ... + c_{n_C}}{e^{j\omega n_C} + c_1 e^{j\omega(n_C-1)} + ... + c_{n_C}} \tag{23}$$

We show that A_m^N decrease like $\xi_2(m) = \frac{\beta_1}{m^2} + \frac{\beta_2}{m^4}$ for $m \geq 1$ where β_1, β_2 are determined with well chosen values of m. We define the *large order* $\mathcal{L}$ to limit the development of (22) by the condition $\xi_2(\mathcal{L}) = \tau$ where τ is a threshold corresponding to 1% of $max\left(A_m^N\right)$. The large order is then given by

$$\mathcal{L} = \mathcal{F}\left[\sqrt{\frac{1}{2\tau}\left(\sqrt{\left(\beta_1^N\right)^2 + 4\beta_2^N\tau} + \beta_1^N\right)}\right] \tag{24}$$

Moreover we show that $\sup_k \left\|\psi_k\left(\hat{\theta}_N^H,\gamma^*\right) - \psi_k^{\mathcal{L}}\left(\hat{\theta}_N^H,\gamma^*\right)\right\| \leq \frac{C}{(\mathcal{L})^2}$ meaning that the bias decreases like $\frac{1}{\mathcal{L}^2}$, ensuring a good convergence of ψ_k. The limited expression of $\psi_k\left(\hat{\theta}_N^H,\gamma^*\right)$ is then yielded by

$$\psi_k^{\mathcal{L}}\left(\hat{\theta}_N^H,\gamma^*\right) = \sum_{m=0}^{\mathcal{L}} A_m^N \varphi_{k-m}(\hat{\theta}_N^H,\gamma^*) \tag{25}$$

Analogous approach can be made for $\partial_\theta\psi_k^T(\hat{\theta}_N^H,\gamma^*)$. Indeed, its limited Taylor's development has the same large order $\mathcal{L}$ and we show that $\sup_k \left\|\partial_\theta\psi_k(\hat{\theta}_N^H,\gamma^*)^T - \partial_\theta\psi_k^{\mathcal{L}}(\hat{\theta}_N^H,\gamma^*)^T\right\|_\infty \leq \frac{C}{\mathcal{L}^2}$. We then get

$$\partial_\theta\psi_k^{\mathcal{L}}(\hat{\theta}_N^H,\gamma^*)^T = C_k\left(\hat{\theta}_N^H,\gamma^*\right) + C_k^T\left(\hat{\theta}_N^H,\gamma^*\right) \tag{26}$$

where the matrix $C_k\left(\hat{\theta}_N^H, \gamma^*\right) \in \mathbb{R}^{d\times d}$ is

$$C_k\left(\hat{\theta}_N^H, \gamma^*\right) = \begin{pmatrix} O_{n_A\times d} \\ ------------- \\ -\sum_{m=0}^{\mathcal{L}}\sum_{l=0}^{\mathcal{L}} A_m^N A_l^N \varphi_{k-1-m-l}^T(\hat{\theta}_N^H, \gamma^*) \\ \cdots \\ \cdots \\ -\sum_{m=0}^{\mathcal{L}}\sum_{l=0}^{\mathcal{L}} A_m^N A_l^N \varphi_{k-n_C-m-l}^T(\hat{\theta}_N^H, \gamma^*) \end{pmatrix} \tag{27}$$

In the following section, proof of the asymptotic convergence in law of $\sqrt{N}\left(\hat{\theta}_N^H - \theta^*\right)$ is considered. This requires the stochastic differentiability and m-dependence approaches.

2.6. Asymptotic convergence in law

For the asymptotic convergence in law of $\sqrt{N}\left(\hat{\theta}_N^H - \theta^*\right)$, let us consider the following technical points related to the signal models of $\varepsilon_k(\theta^*, \gamma^*)$ and $\psi_k(\theta^*, \gamma^*)$.

2.6.1. Signal models

Assume $\tilde{Z}^\infty = \lim_{N\to\infty} \tilde{Z}^N$ the data set and consider $\left(\Omega_j(\theta^*, \gamma^*)\right)_{j\in\nu_1(\theta^*,\gamma^*)}$, $\left(\phi_j(\theta^*, \gamma^*)\right)_{j\in\nu_1(\theta^*,\gamma^*)}$ the NO in $\varepsilon_k(\theta^*, \gamma^*)$ and $\psi_k(\theta^*, \gamma^*)$ respectively. We can write

$$\varepsilon_k(\theta^*, \gamma^*) = \underbrace{\sum_{m\geq 0} \beta_{k,m}(\theta^*, \gamma^*)\, e_{k-m}}_{k\in\nu_2(\theta^*,\gamma^*)} + \underbrace{\sum_{j} \Omega_j(\theta^*, \gamma^*)\, \delta_{k,j}}_{k\in\nu_1(\theta^*,\gamma^*)} \tag{28}$$

$$\psi_k(\theta^*, \gamma^*) = \underbrace{\sum_{m\geq 0} \alpha_{k,m}(\theta^*, \gamma^*)\, e_{k-m}}_{k\in\nu_2(\theta^*,\gamma^*)} + \underbrace{\sum_{j} \phi_j(\theta^*, \gamma^*)\, \delta_{k,j}}_{k\in\nu_1(\theta^*,\gamma^*)} \tag{29}$$

for some filters

$$\{\alpha_{k,m}(\theta^*, \gamma^*), \beta_{k,m}(\theta^*, \gamma^*)\} = f_{k,m}(\theta^*, \gamma^*)$$

Here $\delta_{t,j}$ is the Kronecker function and

H1:

1. $\{e_k\}$ is a sequence of independent rv's with zero mean values and bounded moments of order $4+\delta$, for $\delta > 0$.
2. The family of filters $f_{k,m}(\theta^*, \gamma^*)$, $k = 1, 2, ...$ is uniformly stable for all k, θ^*, γ^* with $f_{k,m}(\theta^*, \gamma^*) < \mu_m$ and $\sum_{m\geq 0} \mu_m < \infty$.

3. Natural outliers $\Omega_j(\theta^*, \gamma^*)$ and $\phi_j(\theta^*, \gamma^*)$ are bounded for all θ^*, γ^* and j, $\sup_{j,\theta^*,\gamma^*} \left|\Omega_j(\theta^*, \gamma^*)\right| = \hat{\Omega}$ and $\sup_{j,\theta^*,\gamma^*} \left|\phi_j(\theta^*, \gamma^*)\right| = \hat{\phi}$.

2.6.2. Stochastic differentiability

In the literature, the standard asymptotic normality results for MLE requires that (4) be twice continuously differentiable, which is not the case here by the presence of the sign function. There exists, however, asymptotic normality results for non-smooth functions and we will hereafter use the one proposed by Newey and McFadden [31] and Andrews [6]. The basic insight of their approaches is that the smoothness condition of (4), $W_N(\theta, \gamma)$ can be replaced by a smoothness of its limit, which in the standard maximum likelihood case corresponds to the expectation $-\overline{E} ln f_H(\varepsilon_k(\theta, \gamma)) = \overline{W}(\theta, \gamma)$, with the requirement that certain remainder terms are small. Hence, the standard differentiability assumption is replaced by a *stochastic differentiability* condition, which can then be used to show that the MLE $\hat{\theta}_N^H$ is asymptotically normal. Recall that the derivative w.r.to θ of ρ_γ^H is $\Psi_k(\theta, \gamma)$. If this function is differentiable in θ, one can establish the asymptotic normality of $\hat{\theta}_N^H$ by expanding $\sqrt{N}\left(\hat{\theta}_N^H - \theta^*\right)$ about θ^* using element by element mean value expansions. This is the standard way of establishing asymptotic normality of the estimator. In a variety of applications, however, $\Psi_k(\theta, \gamma)$ is not differentiable in θ, or not even continuous, due to the appearance of a sign function. In such a case, one can still establish asymptotic normality of the estimator provided $\overline{E}\Psi_k(\theta, \gamma)$ is differentiable in θ. Since the expectation operator is a smoothing operator, $\overline{E}\Psi_k(\theta, \gamma)$ is often differentiable in θ, even though $\Psi_k(\theta, \gamma)$ is not.

2.6.3. m-dependence

Let us consider m a non-negative interger, then a sequence X_ν of random variables is m-dependent if $X_1, X_2, ..., X_s$ is independent of $X_k, X_{k+1}, ...$ provided $k - s > m$ [32] [38]. Here, this approach is applied since the terms in $\partial_\theta W_N(\theta, \gamma)$ are not independent. The purpose is to split the sum in (11) into one part that satisfies a certain independence condition (m-dependence) among its terms and one part that is small. With assumptions H1, the dependence between distant terms will decrease. Thus, let us consider two following lemmas

Lemma 1

Consider the sum of doubly indexed rv's $\{x_{k,N}\}$ such that $S_N = \sum_{k=1}^{N} x_{k,N}$, where $Ex_{k,N} = 0$ and $\{x_{1,N}, ..., x_{s,N}\}$, $\{x_{k,N}, x_{k+1,N}, ..., x_{n,N}\}$ are independent for $k - s > m$. If

$$\lim_{N\to\infty} \sup \sum_{k=1}^{N} Ex_{k,N}^2 < \infty \tag{30}$$

and

$$\lim_{N\to\infty} \sum_{k=1}^{N} E\left|x_{k,N}\right|^{2+\delta} = 0, \ \ \delta > 0, \ \ \text{Lyapunov's condition} \tag{31}$$

, then S_N is asymptotically normal distributed with zero mean and covariance matrix $Q = \lim_{N\to\infty} ES_N S_N^T$. See [32] and [38].

Lemma 2

Let $S_N = Z_{m,N} + X_{m,N}$, $m, N = 1, 2, \ldots$ such that

- $EX_{m,N}^2 \leq C_m$, $\lim_{m\to\infty} C_m = 0$.
- $P(Z_{m,N} \leq z) = F_{m,N}(z)$.

Then $\lim_{m\to\infty}\lim_{N\to\infty} P(Z_{m,N} \leq z) = F(z)$. See [16] and [5].

To prove the asymptotic normality of $\sqrt{N}\left(\hat{\theta}_N^H - \theta^*\right)$, signal models, stochastic differentiability and m-dependence are required. Let us consider the following theorem

theorem 2. *Let $\varepsilon_1(\theta^*, \gamma^*), \ldots, \varepsilon_N(\theta^*, \gamma^*)$ be iid rv's from the pdf f_H with an unknown parameter θ^*, $\theta^* \in \mathcal{D}_c^\theta$ with $\mathcal{D}_c^\theta$ a compactness and $\check{\mathcal{D}}_c^\theta$ interior of $\mathcal{D}_c^\theta$. Then the MLE $\hat{\theta}_N^H$ of θ^* is asymptotically normal*

$$\sqrt{N}\left(\hat{\theta}_N^H - \theta^*\right) \xrightarrow{d} \mathcal{N}\left(0, \mathcal{P}(\theta^*, \gamma^*)\right) \tag{32}$$

where $\mathcal{P}(\theta^*, \gamma^*)$ is the asymptotic covariance matrix given by (21).

In order to do so, all the following assumptions hold. Suppose $W_N(\hat{\theta}_N^H, \gamma^*) \geq \sup_{\theta\in D_M, \gamma^*\in\mathcal{D}_c^\gamma} W_N(\theta, \gamma^*) - o_p\left(N^{-1}\right)$, $\hat{\theta}_N^H \overset{prob}{\to} \theta^*$, and

(i) $W(\theta, \gamma^*)$ is maximized on D_M at θ^*

(ii) θ^* is an interior point of D_M

(iii) $W(\theta, \gamma)$ is twice differentiable at (θ^*, γ^*) with nonsingular second derivative $\overline{\partial^2_{\theta\theta} W}_N(\theta, \gamma)$

(iv) $\sqrt{N}\left(E\partial_\theta W_N(\theta, \gamma^*)\right)_{\hat{\theta}_N^H} \xrightarrow{d} \mathcal{N}\left(0, Q(\theta^*, \gamma^*)\right)$

(v) For any $\delta_N \to 0$, $\sup_{\|\hat{\theta}_N^H - \theta^*\| \leq \delta_N, \gamma^* \to \mathcal{D}_c^\gamma} \left| \frac{\hat{R}_N(\hat{\theta}_N^H, \gamma^*)}{1 + \sqrt{N}\|\hat{\theta}_N^H - \theta^*\|} \right| \overset{prob}{\to} 0$ with the remainder

$$\hat{R}_N(\theta, \gamma^*) = \sqrt{N}\frac{W_N(\theta, \gamma^*) - W_N(\theta^*, \gamma^*) - \left(\partial_\theta W_N(\theta, \gamma^*)\right)_{\theta^*}(\theta - \theta^*) - W(\theta, \gamma^*) + W(\theta^*, \gamma^*)}{\|\theta - \theta^*\|} \tag{33}$$

then $\sqrt{N}\left(\hat{\theta}_N^H - \theta^*\right) \xrightarrow{d} \mathcal{N}\left(0, \mathcal{P}(\theta^*, \gamma^*)\right)$. The proof is given in Appendix A.

3. Experimental results

Experimental results are presented over 16 CO, 15 PD, and 19 HD, left and right feet for different estimation norms. The L_2 norm corresponds to the LSE (least square estimation), L_1 norm to the least sum absolute deviation (LSAD) and L_∞ norm to the supremum norm given by ($\hat{\theta}_N^\infty = \min_\theta \max_t |\varepsilon_t(\theta, \gamma^*)|$). In the Huberian context, a campaign of estimations is carried out in C_γ with $\gamma^* = 1.5$ ([22]) and E_γ with $0.001 \leq \gamma^* \leq 0.2$. For each estimator, comparisons between CO vs PD and HD for left and right feet are given. Table. A.1 shows the means of $\gamma*$, $RMSE$, $FIT(\%)$, $L_2C(\%)$, $L_1C(\%)$ and the total number of parameters $n = n_A + n_C$. The RMSE is the root mean square error between process output and prediction model output. The FIT is given by $100\left(1 - \frac{y-\hat{y}}{y-<y>}\right)$ where y, $\hat{y}$ and $< y >$ are the process output, the prediction model output and the mean of the process output, respectively. L_2C and L_1C are the L_2 and L_1 contributions respectively given by $L_iC = \frac{card\left[\nu_i\left(\hat{\theta}_N^H,\gamma^*\right)\right]}{N}$. These are indicators of the density of NO in the prediction errors. If $L_2C = 40\%$ this means that 40% of prediction errors belong to the interval $[-\gamma^*, \gamma^*]$ and deal with the L_2 norm in the Huberian function. Here, the threshold γ in E_γ was varied among the range [0.001; 0.2] with an incremental step of 0.001 for CO, PD and HD. We focus on the main results in Table. A.1. First, the L_2, L_1 and L_∞ norms give bad results with large RMSE, low FIT and large number of parameters between 40 and 70. The lacks of robustness and degree of freedom (DOF) in these norms lead to an overestimation of the number of parameters n. On the other hand, each FIT presents a low value. In C_γ for $\gamma^* = 1.5$, the number of parameters is reduced with $25 \leq n \leq 32$ but not sufficient for a reduced order ARMA modeling. We can notice a great L_2 contribution, meaning a too large contribution of the L_2 norm, very sensitive to the large NO in the prediction errors.

The Huberian approach in E_γ leads to relevant results. Indeed, this remains in agreement with the formal point of view related to the bias and the new curve in section 2.3: low values of γ involve reduced bias and improve the FIT of the reduced order model. In Corbier and Carmona [15] we showed that the Huberian model order denoted d_{M}^H is such that $d_{\mathrm{M}}^H < d_{\mathrm{M}}^{L_1} < d_{\mathrm{M}}^{L_2}$ since the Huberian function has one DOF and can be tuned from γ, by improving the estimation and reducing the number of parameters for pseudolinear models.

First we notice that $< \gamma^*_{control} > \approx 2 < \gamma^*_{disease} >$, meaning that there are twice more NO in STS-PD and STS-HD than STS-CO. Indeed, for PD and HD, the estimation requires a low value of γ^* involving a large value of the L_1 contribution close to 70%. For CO, $\gamma^* \approx 0.19$ and $L_1C \approx 58\%$. Table. A.2 shows the parameters and variance of each parameter for CO and PD left with $\gamma^* = 0.05$ and $\gamma^* = 0.003$ respectively. For the variance/covariance matrix of these models, the large order $\mathcal{L}$ is equal to 10 ensuring a low computational cost of $C_k\left(\hat{\theta}_N^H, \gamma^*\right)$. Table. A.3 yields the coefficients A_m^N for $m = 0..10$. Figure. A.4 and A.5 show two ARMA models for left CO ($\gamma^* = 0.05$) and left PD ($\gamma^* = 0.003$) respectively with a FIT close to 83%. In Figure. A.5 NO clearly appear in index-times $k = 52$, $k = 113$, $k = 190$ and $k = 247$ with high levels corresponding to the turn around during the walking period. In this phase, the classical estimators are highly disturbed and achieve sometimes the leverage point [22]. We can notice the good behavior of the Huberian reduced order ARMA model during this phase. Equation (34) shows the reduced order ARMA model of left PD for $\gamma^* = 0.003$.

$$\delta t_k = 0,712\delta t_{k-1} + 0,022\delta t_{k-2} + 0,018\delta t_{k-3} + 0,181\delta t_{k-4} + 0,060\delta t_{k-5} + e_k - 0,236e_{k-1} - 0,065e_{k-2} + 0.141e_{k-3} - 0,098e_{k-4} \tag{34}$$

The limited number of ARMA parameters contradicts conclusions in [20] and recently in [1]. These studies showed a stride intervals of normal human walking which exhibit long-range temporal correlations. They presented a highly simplified walking model by reproducing the long-range correlations observed in stride intervals without complex peripheral dynamics. Based on fractal approach they showed an important point of view related to the long-range *memory effect* of human walking. Our new approach shows a *short-range memory effect* for normal and disease human walking. It remains to investigate this *memory effect* and try to interpret in physiological terms the correlations with the CNS.

4. Conclusion

The main purpose of this paper has been to present a reduced order ARMA estimation method based on a robust approach using Huberian function for the neurodegenerative disorder signal modeling. A new approach has been presented to choose the threshold in Huberian function, allowing a best treatment of the natural outliers contained in the signals. The reduced number of parameters is due to a relevant choice of this threshold in a new interval range. Convergence and consistency properties of the robust estimator have been shown including stochastic differentiability and m-dependence approaches. An estimations campaign has been conducted from STS real measurements and it has been shown the relevance to use a Huberian function with DOF to tune its threshold in order to assess a reduced order ARMA model. However, it remains to characterize more appreciably the diseases to differentiate the neurodegenerative disorders. Accordingly, future work will focus on mixed L_p estimator [15] to reduce the number of parameters providing new indicators and will investigate the *memory effect* of human walking.

Appendix A. Proof of the theorem 2

(i): From $Eln f_H(\varepsilon_k(\theta,\gamma^*))$, we can deduce that

$$\theta^* = \underset{\theta\in D_M,\gamma^*\in\mathcal{D}_c^\gamma}{argmax}\left(\frac{-1}{N}\sum_{k=1}^{N}\rho_\gamma^H(\varepsilon_k(\theta,\gamma^*))\right), \text{ as } N\to\infty \tag{A.1}$$

which is equivalent to

$$\underset{\theta\in D_M,\gamma^*\in\mathcal{D}_c^\gamma}{argmin}\left(\frac{1}{N}\sum_{k=1}^{N}\rho_\gamma^H(\varepsilon_k(\theta,\gamma^*))\right), \text{ as } N\to\infty \tag{A.2}$$

Since $W(\theta,\gamma^*) = E\left(\rho_\gamma^H(\varepsilon_k(\theta,\gamma^*))\right)$, then $W(\theta,\gamma^*)$ is maximized on D_M at θ^*.

(ii): The interior condition is equivalent to the assumption $\theta^*\in\mathcal{D}_c^\theta$ where $\check{\mathcal{D}}_c^\theta$ is the interior of D_M.

(iii): Using the stochastic differentiability condition, $E\partial^2_{\xi\xi}\rho_\gamma^H(\varepsilon_k(\theta^*,\gamma^*)) = \overline{\partial^2_{\xi\xi}W}(\theta^*,\gamma^*)$ is invertible as $N\to\infty$.

(iv): Using the mean value theorem, we get

$$\left(E\partial^2_{\xi\xi}W_N(\xi,\gamma^*)\right)_{\tilde{\theta}_N}\left(\hat{\theta}_N^H-\theta^*\right) = \left(E\partial_\theta W_N(\theta,\gamma^*)\right)_{\hat{\theta}_N^H} - \left(E\partial_\theta W_N(\theta,\gamma^*)\right)_{\theta^*} \tag{A.3}$$

with $\hat{\theta}_N^H\le\tilde{\theta}_N\le\theta^*$. For $N\to\infty$, $\tilde{\theta}_N\to\theta^*$, $(E\partial_\theta W_N(\theta,\gamma))_{\theta^*}=0$ and $\lim_{N\to\infty}\left(E\partial^2_{\xi\xi}W_N(\xi,\gamma)\right)_{\tilde{\theta}_N}\to\overline{\partial^2_{\theta\theta}W}(\theta^*,\gamma^*)$. One has

$$\sqrt{N}\left(\hat{\theta}_N^H-\theta^*\right) = \left(\overline{\partial^2_{\theta\theta}W}(\theta^*,\gamma^*)\right)^{-1}\sqrt{N}\left(E\partial_\theta W_N(\theta,\gamma^*)\right)_{\hat{\theta}_N^H} \tag{A.4}$$

The asymptotic normality of $\sqrt{N}\left(\hat{\theta}_N^H-\theta^*\right)$ only depends on the asymptotic normality of $\sqrt{N}\left(E\partial_\theta W_N(\theta,\gamma^*)\right)_{\hat{\theta}_N^H}$. Let us denote $\partial_\theta W_N(\theta,\gamma) = \frac{-1}{N}\sum_{k\in\nu_2(\theta,\gamma)}\psi_k(\theta,\gamma)\varepsilon_k(\theta,\gamma) - \frac{\gamma}{N}\sum_{k\in\nu_1(\theta,\gamma)}\psi_k(\theta,\gamma)s_k(\theta,\gamma) = \frac{1}{N}\sum_{k=1}^{N}\check{\Psi}_k(\theta,\gamma)$. Therefore,

$$-\sqrt{N}\left(E\partial_\theta W_N(\theta,\gamma^*)\right)_{\hat{\theta}_N^H} = \sqrt{N}\left\{\frac{1}{N}\sum_{k=1}^{N}\left[\check{\Psi}_k\left(\hat{\theta}_N^H,\gamma^*\right)-E\check{\Psi}_k\left(\hat{\theta}_N^H,\gamma^*\right)\right]\right\} - \sqrt{N}\frac{1}{N}\sum_{k=1}^{N}\check{\Psi}_k\left(\hat{\theta}_N^H,\gamma^*\right)$$

that is

$$= \frac{1}{\sqrt{N}}\sum_{k=1}^{N}\left(\check{\Psi}_k\left(\hat{\theta}_N^H,\gamma^*\right)-E\check{\Psi}_k\left(\hat{\theta}_N^H,\gamma^*\right)\right) - \frac{1}{\sqrt{N}}\sum_{k=1}^{N}\check{\Psi}_k\left(\hat{\theta}_N^H,\gamma^*\right) \tag{A.5}$$

Let us denote $S_N(\theta,\gamma^*) = \frac{1}{\sqrt{N}}\sum_{k=1}^{N}\left(\check{\Psi}_k(\theta,\gamma^*)-E\check{\Psi}_k(\theta,\gamma^*)\right)$, then

$$-\sqrt{N}\left(E\partial_\theta W_N(\theta,\gamma^*)\right)_{\hat{\theta}_N^H} = \left(S_N\left(\hat{\theta}_N^H,\gamma^*\right)-S_N(\theta^*,\gamma^*)\right)+S_N(\theta^*,\gamma^*)-\frac{1}{\sqrt{N}}\sum_{k=1}^{N}\check{\Psi}_k\left(\hat{\theta}_N^H,\gamma^*\right) \tag{A.6}$$

Since $\frac{1}{N}\sum_{k=1}^{N}\check{\Psi}_k\left(\hat{\theta}_N^H,\gamma^*\right)=0$, the third term on the right hand side of (A.6) is $o(1)$. Its first term is $o(1)$ provided $\{S_N(\bullet,\gamma^*), N\ge 1\}$ is stochastically equicontinuous and $\hat{\theta}_N^H\overset{prob}{\to}\theta^*$. This follows because given any $\alpha>0$ and

$\beta > 0$, there exists a $\delta > 0$ such that for $\Delta S\left(\hat{\theta}_N^H, \theta^*, \gamma^*\right) = S_N(\hat{\theta}_N^H, \gamma^*) - S_N(\theta^*, \gamma^*)$

$$\overline{\lim_{N\to\infty}} P\left(\left|\Delta S\left(\hat{\theta}_N^H, \theta^*, \gamma^*\right)\right| > \alpha\right) \leq$$

$$\overline{\lim_{N\to\infty}} P\left(\left|\Delta S\left(\hat{\theta}_N^H, \theta^*, \gamma^*\right)\right|, \left\|\rho_\gamma^H(\varepsilon_k(\hat{\theta}_N^H), \gamma^*) - \rho_\gamma^H(\varepsilon_k(\theta^*), \gamma^*)\right\| \leq \delta\right) + \overline{\lim_{N\to\infty}} P\left(\left\|\rho_\gamma^H(\varepsilon_k(\hat{\theta}_N^H), \gamma^*) - \rho_\gamma^H(\varepsilon_k(\theta^*), \gamma^*)\right\| > \delta\right) \tag{A.7}$$

$$\leq \overline{\lim_{N\to\infty}} P\left(\sup_{\theta\in D_M, \gamma^*\in\mathcal{D}_c^\gamma} |S_N(\theta, , \gamma^*) - S_N(\theta^*, \gamma^*)| > \alpha\right) < \beta \tag{A.8}$$

where the second inequality uses $\hat{\theta}_N^H \overset{prob}{\to} \theta^*$ and the third uses the stochastic equicontinuity. Accordingly, for a given threshold γ^*, this shows that for N tends to infinity, we have in law

$$\mathcal{L}\left(\sqrt{N}\left(\hat{\theta}_N^H - \theta^*\right)\right) \underset{N\to\infty}{\sim} \mathcal{L}(S_N(\theta^*, \gamma^*)) \tag{A.9}$$

with

$$S_N(\theta^*, \gamma^*) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \left(\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_k(\theta, \gamma^*) - E\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_k(\theta, \gamma^*)\right)_{\theta^*} \tag{A.10}$$

The purpose is to prove that $S_N(\theta^*, \gamma^*)$ is a normal asymptotic distribution. For this, we show that the terms of $S_N(\theta^*, \gamma^*)$ are independent. As described above, we use the m-dependence approach to show the asymptotic normal behavior of $S_N(\theta^*, \gamma^*)$. Let us consider the following short expressions: $\varepsilon_{v_i,k}(\theta^*, \gamma^*) = \varepsilon_{i,k}^*$, $f_{t,k}(\theta^*, \gamma^*) = f_{t,k}^*$. We split $\varepsilon_{2,t}^*$ and $\psi_{2,t}^*$ into one part that satisfies m-dependence conditions among its terms and one part that is small. We then have

$$\varepsilon_{2,t}^* = \varepsilon_{2,t}^{*,m} + \tilde{\varepsilon}_{2,t}^{*,m} + \varepsilon_{1,t}^* = \sum_{k=0}^{m} \beta_{t,k}^* e_{t-k} + \sum_{k=m+1}^{\infty} \beta_{t,k}^* e_{t-k} + \sum_j \Omega_j^* \delta_{t,j}^K \tag{A.11}$$

where m is an integer with $\Omega_j^* = \Omega_j(\theta^*, \gamma^*)$ and $\delta_{t,j}^K$ is the Kronecker's function. Analogously, we have

$$\psi_{2,t}^* = \psi_{2,t}^{*,m} + \tilde{\psi}_{2,t}^{*,m} + \psi_{1,t}^* = \sum_{k=0}^{m} \alpha_{t,k}^* e_{t-k} + \sum_{k=m+1}^{\infty} \alpha_{t,k}^* e_{t-k} + \sum_j \phi_j^* \delta_{t,j}^K \tag{A.12}$$

$S_N(\theta^*, \gamma^*)$ can be written as $S_N(\theta^*, \gamma^*) = Z_{m,N}(\theta^*, \gamma^*) + X_{m,N}(\theta^*, \gamma^*)$ with

$$Z_{m,N}(\theta^*, \gamma^*) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} \left(\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*)) - E\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))\right)_{\theta^*} \tag{A.13}$$

$$X_{m,N}(\theta^*, \gamma^*) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} \frac{d}{d\theta}\left[\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*)) - \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))\right]_{\theta^*} - E\frac{d}{d\theta}\left[\rho_\gamma^H(\varepsilon_t(\theta, \gamma^*)) - \rho_\gamma^H(\varepsilon_t^m(\theta, \gamma^*))\right]_{\theta^*} \tag{A.14}$$

Part1:

From (A.13) in $Z_{m,N}(\theta^*,\gamma^*)$ and using the Lyapunov's condition, we obtain

$$E\left|\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*)) - E\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))\right|^{\delta+2} \leq 2^{\delta+1}E\left(\left|\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))\right|^{\delta+2} + E\left|\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))\right|^{\delta+2}\right) \quad (A.15)$$

$$\leq 2^{\delta+2}E\left|\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))\right|^{\delta+2} \quad (A.16)$$

with

$$\left|\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))\right|^{\delta+2} \leq (\left|\psi_{2,t}^{*,m}\right|\left|\varepsilon_{2,t}^{*,m}\right| + \gamma^*\left|\psi_{1,t}^{*,m}\right|)^{\delta+2} \leq 2^{\delta+1}(\left|\psi_{2,t}^{*,m}\right|^{\delta+2}\left|\varepsilon_{2,t}^{*,m}\right|^{\delta+2} + (\gamma^*)^{\delta+2}\left|\psi_{1,t}^{*,m}\right|^{\delta+2}) \quad (A.17)$$

We deduce

$$2^{\delta+2}E\left|\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))\right|^{\delta+2} \leq 2^{2\delta+3}E\left|\psi_{2,t}^{*,m}\right|^{\delta+2}\left|\varepsilon_{2,t}^{*,m}\right|^{\delta+2} + 2^{2\delta+3}(\gamma^*)^{\delta+2}E\left|\psi_{1,t}^{*,m}\right|^{\delta+2} \quad (A.18)$$

Using Schwarz's inequality

$$2^{\delta+2}E\left|\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))\right|^{\delta+2} \leq 2^{2\delta+3}\left(E\left|\psi_{2,t}^{*,m}\right|^{2\delta+4}E\left|\varepsilon_{2,t}^{*,m}\right|^{(2\delta+4)}\right)^{\frac{1}{2}} + 2^{2\delta+3}(\gamma^*)^{\delta+2}\left(E\left|\psi_{1,t}^{*,m}\right|^{2\delta+4}\right)^{\frac{1}{2}} \quad (A.19)$$

The first and second terms on the right hand side of (A.19) are respectively denoted A^* and B^*.

- For A^*: in v_2, for all t and θ^*, $\left|\varepsilon_{2,t}^{*,m}\right| \leq \gamma^*$. Therefore

$$E\left|\psi_{2,t}^{*,m}\right|^{2\delta+4} \leq 2^{2\delta+3}E\left|e_{t-k}\right|^{2\delta+4}\left(\sum_{k=0}^{m}\mu_k\right)^{2\delta+4} \quad (A.20)$$

  From **H1**, we have $E\left|\psi_{2,t}^{*,m}\right|^{2\delta+4} \leq C^*$ and $A^* \leq C^*$.

- For B^*: from **H1** we get $\sup_{t,\theta^*,\gamma^*}\left|\varepsilon_{1,t}^{*,m}\right| = \hat{\Omega}$ and $E\left|\psi_{1,t}^{*,m}\right|^{2\delta+4}$ are bounded. Accordingly, $B^* \leq C^*$.

Inserting $\left(\frac{1}{\sqrt{N}}\right)^{\delta+2}$, we finally obtain for all γ^*

$$E\left|\frac{1}{\sqrt{N}}\left[\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*)) - E\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))\right]\right|_{\theta^*}^{\delta+2} \leq \frac{C}{N^{1+\frac{\delta}{2}}} \quad (A.21)$$

Then

$$\lim_{N\to\infty}\sum_{t=1}^{N}E\left|\frac{1}{\sqrt{N}}\left[\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*)) - E\frac{d}{d\theta}\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))\right]\right|_{\theta^*}^{\delta+2} \leq \lim_{N\to\infty}\frac{C}{N^{\frac{\delta}{2}}} \to 0 \quad (A.22)$$

Expression (A.22) proves (30) and (31) in lemma 1 (section 2.6.3) with $Q_m(\theta^*, \gamma) = \lim_{N\to\infty} EZ_{m,N}(\theta^*, \gamma^*)Z^T_{m,N}(\theta^*, \gamma^*)$.

Part2:

In $X_{m,N}(\theta^*, \gamma^*)$, we can write

$$\frac{d\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{d\theta} - \frac{d\rho^H_\gamma(\varepsilon^m_t(\theta,\gamma^*))}{d\theta} = \frac{\partial\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{\partial\varepsilon_t(\theta,\gamma^*)}\frac{\partial\varepsilon_t(\theta,\gamma^*)}{\partial\theta} - \frac{\partial\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{\partial\varepsilon_t(\theta,\gamma^*)}\frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta}$$

$$+\frac{\partial\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{\partial\varepsilon_t(\theta,\gamma^*)}\frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta} - \frac{\partial\rho^H_\gamma(\varepsilon^m_t(\theta,\gamma^*))}{\partial\varepsilon^m_t(\theta,\gamma^*)}\frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta} \tag{A.23}$$

Therefore

$$\frac{d\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{d\theta} - \frac{d\rho^H_\gamma(\varepsilon^m_t(\theta,\gamma^*))}{d\theta} = \frac{\partial\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{\partial\varepsilon_t(\theta,\gamma^*)}\left(\frac{\partial\varepsilon_t(\theta,\gamma^*)}{\partial\theta} - \frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta}\right)$$

$$+\left(\frac{\partial\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{\partial\varepsilon_t(\theta,\gamma^*)} - \frac{\partial\rho^H_\gamma(\varepsilon^m_t(\theta,\gamma^*))}{\partial\varepsilon^m_t(\theta,\gamma^*)}\right)\frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta} \tag{A.24}$$

Using mean value theorem, we get

$$\frac{\partial\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{\partial\varepsilon_t(\theta,\gamma^*)} - \frac{\partial\rho^H_\gamma(\varepsilon^m_t(\theta,\gamma^*))}{\partial\varepsilon^m_t(\theta,\gamma^*)} = \frac{\partial^2\rho^H_\gamma(\xi_t(\theta,\gamma^*))}{\partial\xi_t(\theta,\gamma^*)^2}(\varepsilon_t(\theta,\gamma^*) - \varepsilon^m_t(\theta,\gamma^*)) \tag{A.25}$$

Hence

$$\left|\frac{d\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{d\theta} - \frac{d\rho^H_\gamma(\varepsilon^m_t(\theta,\gamma^*))}{d\theta}\right| \le \left|\frac{\partial\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{\partial\varepsilon_t(\theta,\gamma^*)}\right|\left|\frac{\partial\varepsilon_t(\theta,\gamma^*)}{\partial\theta} - \frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta}\right|$$

$$+\left|\frac{\partial^2\rho^H_\gamma(\xi_t(\theta,\gamma^*))}{\partial\xi_t(\theta,\gamma^*)^2}\right|\left|\varepsilon_t(\theta,\gamma^*) - \varepsilon^m_t(\theta,\gamma^*)\right|\left|\frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta}\right| \tag{A.26}$$

From regularity conditions C1 in (see [25]) given by

- $\left\|\frac{\partial\rho(\varepsilon)}{\partial\varepsilon}\right\| \le C\,|\varepsilon|,\ \theta \in D_{\mathrm{M}}$, all t.
- $\left\|\frac{\partial\rho(\varepsilon)}{\partial\theta}\right\| \le C\,|\varepsilon|^2,\ \theta \in D_{\mathrm{M}}$, all t.
- $\left\|\frac{\partial^2\rho(\varepsilon)}{\partial\varepsilon^2}\right\| \le C$.

We then have

$$\left|\frac{d\rho^H_\gamma(\varepsilon_t(\theta,\gamma^*))}{d\theta} - \frac{d\rho^H_\gamma(\varepsilon^m_t(\theta,\gamma^*))}{d\theta}\right| \le C\,|\varepsilon_t(\theta,\gamma^*)|\left|\frac{\partial\varepsilon_t(\theta,\gamma^*)}{\partial\theta} - \frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta}\right|$$

$$+C\left|\varepsilon_t(\theta,\gamma^*) - \varepsilon^m_t(\theta,\gamma^*)\right|\left|\frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta}\right| \tag{A.27}$$

In detail

$$\left|\frac{\partial\varepsilon_t(\theta,\gamma^*)}{\partial\theta} - \frac{\partial\varepsilon^m_t(\theta,\gamma^*)}{\partial\theta}\right| = \left|\tilde{\psi}^m_t(\theta,\gamma^*)\right| \le \sum_{k=m+1}^{\infty}\left|\alpha^*_{t,k}\right||e_{t-k}| \le \sum_{k=m+1}^{\infty}\mu_k\,|e_{t-k}| \tag{A.28}$$

and

$$\left|\varepsilon_t(\theta,\gamma^*) - \varepsilon_t^m(\theta,\gamma^*)\right| = \left|\tilde{\varepsilon}_t^m(\theta,\gamma^*)\right| \leq \sum_{k=m+1}^{\infty} \left|\beta_{t,k}^*\right| |e_{t-k}| \leq \sum_{k=m+1}^{\infty} \mu_k |e_{t-k}| \tag{A.29}$$

Expression (A.27) becomes

$$\left|\frac{d\rho_\gamma^H(\varepsilon_t(\theta,\gamma^*))}{d\theta} - \frac{d\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))}{d\theta}\right| \leq C\left(\sum_{k=m+1}^{\infty} \mu_k |e_{t-k}|\right)\left(|\varepsilon_t(\theta,\gamma^*)| + \left|\psi_t^m(\theta,\gamma^*)\right|\right) \tag{A.30}$$

Moreover

$$|\varepsilon_t(\theta,\gamma^*)| + \left|\psi_t^m(\theta,\gamma^*)\right| \leq 2\sum_{k=0}^{m} \mu_k |e_{t-k}| + \sum_{k=m+1}^{\infty} \mu_k |e_{t-k}| \tag{A.31}$$

Therefore

$$\left|\frac{d\rho_\gamma^H(\varepsilon_t(\theta,\gamma^*))}{d\theta} - \frac{d\rho_\gamma^H(\varepsilon_t^m(\theta,\gamma^*))}{d\theta}\right| \leq \alpha_t + \beta_t \tag{A.32}$$

with

$$\alpha_t = 2C\left(\sum_{k=0}^{m}\mu_k\right)\left(\sum_{k=m+1}^{\infty} \mu_k |e_{t-k}|\right) \tag{A.33}$$

and

$$\beta_t = C\left(\sum_{k=m+1}^{\infty} \mu_k |e_{t-k}|\right)^2 \tag{A.34}$$

Therefore

$$X_{m,N}(\theta^*,\gamma^*) \leq \underbrace{\frac{1}{\sqrt{N}}\sum_{t=1}^{N}(\alpha_t - E\alpha_t)}_{X_{m,N}^{\alpha}(\theta^*,\gamma^*)} + \underbrace{\frac{1}{\sqrt{N}}\sum_{t=1}^{N}(\beta_t - E\beta_t)}_{X_{m,N}^{\beta}(\theta^*,\gamma^*)} \tag{A.35}$$

Each term on the right hand side of (A.35) verifies the corollary of the lemma 2B.1 in [27](p.57). Hence, as $m \to \infty$

$$E\left(X_{m,N}^{\alpha}(\theta^*,\gamma^*)\right)^2 \leq K\left(\sum_{k=0}^{m}\mu_k\right)\left(\sum_{k=m+1}^{\infty}\mu_k\right) \to 0 \tag{A.36}$$

$$E\left(X_{m,N}^{\beta}(\theta^*,\gamma^*)\right)^2 \leq K\left(\sum_{k=m+1}^{\infty}\mu_k\right)^2 \to 0 \tag{A.37}$$

Hence, $Z_{m,N}(\theta^*,\gamma^*) \in \mathcal{A}s\mathcal{N}(0, Q_m(\theta^*,\gamma^*))$ and $S_N(\theta^*,\gamma^*) \in \mathcal{A}s\mathcal{N}(0, Q(\theta^*,\gamma^*))$ with $Q(\theta^*,\gamma^*) = \lim_{m\to\infty} Q_m(\theta^*,\gamma^*)$. Which proves the point (*iv*) of the Theorem 2.

(v): Expanding $W(\theta,\gamma^*)$ into Taylor series around θ^*, we get

$$W(\hat{\theta}_N^H,\gamma^*) = W(\theta^*,\gamma^*) + \frac{1}{2}\left(\hat{\theta}_N^H - \theta^*\right)^T \overline{\partial_{\theta\theta}^2 W}(\theta^*,\gamma^*)\left(\hat{\theta}_N^H - \theta^*\right) + o\left(\left\|\hat{\theta}_N^H - \theta^*\right\|^2\right) \tag{A.38}$$

Since $\overline{\partial^2_{\theta\theta}W}(\theta^*, \gamma^*)$ is positive definite and nonsingular, there exists $C > 0$ and a neighborhood of θ^* such that

$$\frac{1}{2}\left(\hat{\theta}^H_N - \theta^*\right)^T \overline{\partial^2_{\theta\theta}W}(\theta^*, \gamma^*)\left(\hat{\theta}^H_N - \theta^*\right) + o\left(\left\|\hat{\theta}^H_N - \theta^*\right\|^2\right) \leq C\left\|\hat{\theta}^H_N - \theta^*\right\|^2 \tag{A.39}$$

we obtain $W(\hat{\theta}^H_N, \gamma^*) \leq W(\theta^*, \gamma^*) + C\left\|\hat{\theta}^H_N - \theta^*\right\|^2$. Moreover

$$W_N(\hat{\theta}^H_N, \gamma^*) - W_N(\theta^*, \gamma^*) + o\left(\frac{1}{N}\right) = W(\hat{\theta}^H_N, \gamma^*) - W(\theta^*, \gamma^*) + \left(\partial_\xi W_N(\xi, \gamma^*)\right)^T_{\theta^*}\left(\hat{\theta}^H_N - \theta^*\right) + \left\|\hat{\theta}^H_N - \theta^*\right\|\hat{R}_N\left(\hat{\theta}^H_N, \gamma^*\right) + o\left(\frac{1}{N}\right) \tag{A.40}$$

Therefore

$$\begin{aligned} W_N(\hat{\theta}^H_N, \gamma^*) - W_N(\theta^*, \gamma^*) &\leq C\left\|\hat{\theta}^H_N - \theta^*\right\|^2 + \left\|\partial_\xi W_N(\xi, \gamma^*)\right\|_{\theta^*}\left\|\hat{\theta}^H_N - \theta^*\right\| \\ &+ \left\|\hat{\theta}^H_N - \theta^*\right\|\left(1 + \sqrt{N}\left\|\hat{\theta}^H_N - \theta^*\right\|\right) o\left(\frac{1}{\sqrt{N}}\right) + o\left(\frac{1}{N}\right) \end{aligned} \tag{A.41}$$

Since $\left\|\partial_\xi W_N(\xi, \gamma^*)\right\|_{\theta^*} \to 0$ as $N \to \infty$, then

$$W_N(\hat{\theta}^H_N, \gamma^*) - W_N(\theta^*, \gamma^*) \leq (C + o(1))\left\|\hat{\theta}^H_N - \theta^*\right\|^2 + o\left(\frac{1}{\sqrt{N}}\right)\left\|\hat{\theta}^H_N - \theta^*\right\| + o\left(\frac{1}{N}\right) \tag{A.42}$$

The remainder $\hat{R}_N\left(\hat{\theta}^H_N, \gamma^*\right)$ can be written as

$$\hat{R}_N\left(\hat{\theta}^H_N, \gamma^*\right) \leq \sqrt{N}\left\|\hat{\theta}^H_N - \theta^*\right\|(K + o(1)) \tag{A.43}$$

then

$$\frac{\hat{R}_N\left(\hat{\theta}^H_N, \gamma^*\right)}{1 + \sqrt{N}\left\|\hat{\theta}^H_N - \theta^*\right\|} \leq \frac{\sqrt{N}\left\|\hat{\theta}^H_N - \theta^*\right\|(K + o(1))}{1 + \sqrt{N}\left\|\hat{\theta}^H_N - \theta^*\right\|} \tag{A.44}$$

Since $\sqrt{N}\left\|\hat{\theta}^H_N - \theta^*\right\| \overset{prob}{\to} 0$ then

$$\sup_{\left\|\hat{\theta}^H_N - \theta^*\right\| \leq \delta_N, \gamma^* \to \mathcal{D}^\gamma_c}\left|\frac{\hat{R}_N\left(\hat{\theta}^H_N, \gamma^*\right)}{1 + \sqrt{N}\left\|\hat{\theta}^H_N - \theta^*\right\|}\right| \overset{prob}{\to} 0 \tag{A.45}$$

which prove the point (v) and finally the theorem 2.

[1] J. Ahn, N. Hogan, Long-range correlations in stride intervals may emerge from non-chaotic walking dynamics, PLOS ONE, 8(9), 2013, pp. 2-10.

[2] M.S. Aslam, M.A.Z. Raja, A new adaptive strategy to improve online secondary path modeling in active noise control system using fractional signal processing approach, Signal Processing 107, 2015, pp. 433-443.

[3] A. Al-Smadi, A least-squares based algorithm for identification of non-gaussian ARMA models, Circuits Systems Signal Processing 26(5), 2007, pp. 715-731.

[4] K.E. Al-Qawasmi, A.M. Al-Smadi, A. Al-Hamami, A robust ARX and ARMA model order estimation via pivot-neighbors comparisons,

Recent Patents on Computing Science. 3, 2010, pp. 33-38.

[5] T.W. Anderson, On asymptotic distributions of estimated parameters of stochastic difference equation, Ann. Math. Stat., 30, 1959, pp. 676-687.

[6] D.W.K. Andrews, Large sample estimation and hypothesis testing, Handbook of Econometric, Elsevier Science 4, 1994, pp. 2248-2292.

[7] V. Barnett, T. Lewis, Outliers in statistical data, 3rd ed. Wiley, 1998, New York.

[8] P.M. Broersen, The quality of models for ARMA processes, IEEE Trans. Signal Process. 46(6), 1998, pp. 1749-1752.

[9] Y. Chakhchoukh, A new robust estimation method for ARMA Models, IEEE Tran on Sig Pro. 58(7), 2010, pp. 3512-3522.

[10] N.I. Chaudhary, M.A.Z. Raja, J.A. Khan, M.S. Aslam, Identification of input nonlinear control autoregressive systems using fractional signal processing approach, Hindawi Publishing Corporation, 2013. Article ID 467276. DOI:10.1155/2013/467276.

[11] N.I. Chaudhary, M.A.Z. Raja, Identification of Hammerstein nonlinear ARMAX systems using nonlinear adaptive algorithms, Nonlinear Dynamics, 2014. DOI:10.1007/s11071-014-1748-8.

[12] C. Corbier, Contribution a l'estimation robuste de modeles dynamiques: application a la commande de systemes dynamiques complexes, Arts et Metiers ParisTech, Thesis, Nov, 29, 2012.

[13] C. Corbier, M. El Badaoui, ARMA Modeling and Nonparametric Probability Density Function of Gait Signal Using $L_2 - L_1$ estimator in Patients with Neuro-Degenerative Disease, Computer Methods in Biomechanics and Biomedical Engineering, Taylor and Francis.... 39th Congress of the Societe de Biomecanique, Valenciennes, France, 27-29 August, 2014, 17(1), pp. 178-179.

[14] C. Corbier, J-C. Carmona, Extension of the Tuning Constant in the Hubers Function for Robust Modeling of Piezoelectric Systems, International Journal of Adaptive Control and Signal Processing, August 2014. DOI:10.1002/acs.2517.

[15] C. Corbier, J-C. Carmona Mixed L_p-estimators Variety for Model Order Reduction in Control Oriented System Identification, Hindawi Publishing Corporation Mathematical Problems in Engineering, July 2014. ID 349070.

[16] P.H. Diananda, Some probability limit theorems with statistical applications, Proc. Cambridge Philos. Soc., 49, 1953, pp. 239-245.

[17] A.J. Fox, Outliers in time series, Journal of the Royal Society, 34(3), 1972, 350-363.

[18] A.S. Hadi, A.H.M. Rahmatullah Imon, M. Werner, Detection of outliers, John Wiley and Sons, Inc. Wires Comp Stat , 1, 2009, 57-70.

[19] F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, W.A. Stahel, Robust statistics: the approach based on influence function, John Wiley and Sons. New York, 1985.

[20] J.M. Hausdorff, C.K. Peng, Z. Ladin, J.Y. Wei, A.L. Goldberger, Is walking a random walk? Evidence for long-range correlations in stride interval of human gait, Journal of Applied Physiology, 78(1), 1995, pp. 349-358.

[21] J.M. Hausdorff, M.E. Cudkowicz, R. Firtion, J.Y. Wei, A.L. Goldberger, Gait variability and basal ganglia disorders:stride-to-stride variations of gait cycle timing in Parkinson's disease and Huntington's disease, Movement disorders, 13(3), 1998, 428-437.

[22] P.J. Huber, E.M. Ronchetti, Robust statistics. 2th (Eds), New York, NY: John Wiley and Sons, 2009.

[23] T. Iluz, E. Gazit, T. Herman, E. Sprecher, M. Brozgol, N. Giladi, A. Mirelman, J.M. Hausdorff, Automated detection of missteps during community ambulation in patients with Parkinson's disease: a new approach for quantifying fall risk in the community setting, Journal of NeuroEngineering and Rehabilitation, 2014, 11-48.

[24] Y-C. Lin, M. Gfoehler, M.G. Pandy, Quantitative evaluation of the major determinants of human gait. J. Biomech, 47, 2014, 1324-1331.

[25] L. Ljung, Convergence analysis of parametric identification methods, IEEE Trans on Automatic Control, vol. AC-23, no.5, 1978, pp. 770-783.

[26] L. Ljung, P.E. Caines, Asymptotic Normality of prediction error estimators for approximate systems models, Stochastics, 3, 1979, 29-46.

[27] L. Ljung, System identification: theory for the user, Prentice Hall PTR. New York, 1999.

[28] S. Maiz, M. El Badaoui, F. Bonnardot, C.Serviere, New order cyclostationary analysis and application to the detection and characterization of a runner's fatigue. Signal Processing 102, 2014, 188-200.

[29] A.E. Martin, J.P. Schmiedeler, Predicting human walking gaits with a simple planar model, J. Biomech, 47, 2014, 1416-1421.

[30] N. Muler, D. Pena, V.J. Yohai, Robust estimation for ARMA models, The Annals of Statistics, 37(2), 2009, pp. 816840.

[31] W.K. Newey and D.L. McFadden, Large sample estimation and hypothesis testing, Handbook of Econometric, Elsevier Science 4, 1994, pp. 2113-2247.

[32] S. Orey, A central limit theorem for m-dependent random variables, Duke Math.J., 25, 1958, pp. 543-546.

[33] M.A.Z. Raja, N.I. Chaudhary, Two-stage fractional least mean square identification algorithm for parameter estimation of CARMA systems, Signal Processing, 107, 2015, pp. 327-339. DOI:10.1016/j.sigpro.2014.06.015.

[34] C. Ran, Z. Deng, Self-tuning distributed measurement fusion Kalman estimator for multi-channel ARMA signal, Signal Processing, 91, 2011, pp. 2028-2041.

[35] L. Ren, R.K. Jones, D. Howard, Predictive modeling of human walking over a complete gait cycle, J.Neurol., Neurosurgery, Psychiatry, 79, 2007, 874-880.

[36] R.T. Roemmich, P.R., Zeilman, M.S.O., Vaillancourt, C.J., Hass, Gait variability magnitude but not structure is altered in essential tremor, J. Biomech, 46(15), 2013, 2682-2687.

[37] A.H. Ropper, R.H., Brown, Adams and Victor's principles of neurology. 8th (Eds), McGraw-Hill, New York, NY, 2005.

[38] B. Rosen, On a central limit theorem for sums of dependent random variables, Z., Wahrsch verw. Geb., 7, 1967, pp. 48-82.

[39] N. Scafetta, D., Marchi, B.J., West, Understanding the complexity of human gait dynamics, Chaos, 19(2), 2009, 026108.

[40] S. Sen Roy, S., Guria, Estimation of regression parameters in the presence of outliers in the response, Statistics, 25(5), 2009, 1000-1009.

[41] S.M. Shah, R. Samar, M.A.Z. Raja, J.A. Chambers, Fractional normalized filtered-error least mean squares algorithm for application in active noise control systems, Electronics Letters 50 (14), 973-975, 2014. DOI: 10.1049/el.2014.1275.

[42] J. Sian, M. Gerlach, M.B.H. Youdim, P., Riederer, Parkinson's disease:a major hypokinetic basal ganglia disorder, J.Neural Transmission, 106(5-6), 1999, 443-476.

[43] Y. Wu, S., Krishnan, Statistical analysis of gait maturation in children using nonparametric probability density function. Entropy, 15, 2013, 753-766.

[44] F. Zacharia, M. El Badaoui, S. Maiz, F. Guillet, M. Khalil, K. Khalil, M. Halimi, Walking analysis: Empirical relation between kutosis and degree of cyclostationary, 2nd International Conference on Advances in Biomedical Engineering IEEE, 11-13 Sept, 2013, pp. 93-96.

[45] K. Zhu and S. Ling , LADE-based inference for ARMA models with unspecified and heavy-tailed heteroscedastic noises, Chinese Academy of Science, Hong Kong University of Science and Technology, MPRA Paper No. 59099, 2014, pp. 2-33.

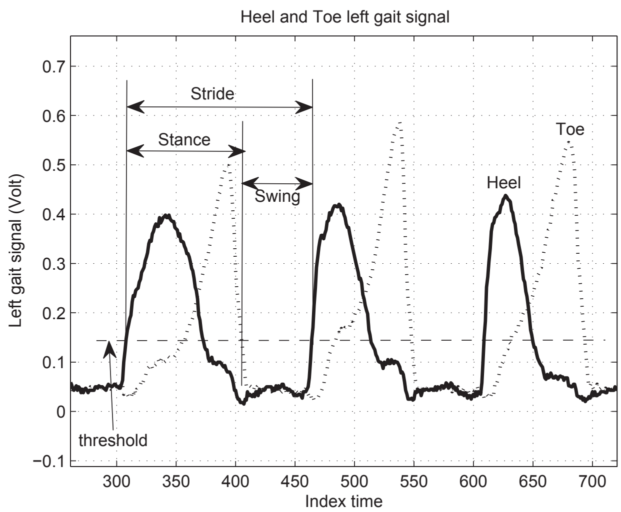

Figure A.1: Example of gait signals from heel and toe force sensors underneath the left foot. The threshold allows to compute the time-signals δt_k such as the stride, swing and stance.

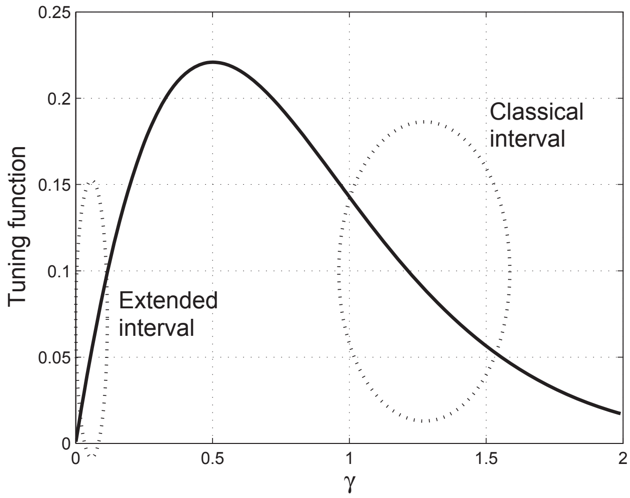

Figure A.2: Tuning function with two main intervals. The classical interval $\gamma \in [1, 1.5]$ and the extended interval $\gamma \in [0.001, 0.2]$.

Table A.1: Means of $\gamma*$, $RMSE$, $FIT(\%)$, $L_2C(\%)$, $L_1C(\%)$ and the total number of parameters $n = n_A + n_C$ over 16 CO 15 PD and 19 HD (left and right feet) for different estimation norms. L_2 is the LSE, L_1 is the LSAD, L_∞ is the supremum norm. C_γ is the classical interval in the Huber's context with $< \gamma* >= 1.5$. E_γ is the extended interval in the Huber's context with low values of $\gamma*$.

			CO left						**PD left**			
Estimator	$\gamma*$	$RMSE$	FIT	L_2C	L_1C	n	$\gamma*$	$RMSE$	FIT	L_2C	L_1C	n
L_2	–	11.2	10	100	0	70	–	13	9	100	0	70
L_1	–	4.3	42	0	100	41	–	5.2	38	0	100	46
L_∞	–	4.2	25	–	–	45	–	5.3	26	–	–	56
Huber in C_γ	1.5	2.4	42	95	5	25	1.5	3.1	31	96	4	28
Huber in E_γ	0.17	0.09	92	41	59	9	0.09	0.34	78	30	70	9
			CO right						**PD right**			
L_2	–	10.2	9	100	0	70	–	13	9	100	0	70
L_1	–	5.3	44	0	100	39	–	6.2	35	0	100	46
L_∞	–	3.2	26	–	–	46	–	5.5	28	–	–	54
Huber in C_γ	1.5	2.3	44	96	4	27	1.5	3.3	31	96	4	30
Huber in E_γ	0.18	0.08	92	43	57	9	0.09	0.29	78	32	68	9
			CO left						**HD left**			
L_2	–	11.2	10	100	0	70	–	8	17	100	0	70
L_1	–	4.3	42	0	100	41	–	4.1	36	0	100	44
L_∞	–	4.2	25	–	–	45	–	6.3	24	–	–	54
Huber in C_γ	1.5	2.4	42	95	5	25	1.5	3.2	32	96	4	31
Huber in E_γ	0.17	0.09	92	41	59	9	0.08	0.28	78	29	71	9
			CO right						**HD right**			
L_2	–	10.2	9	100	0	70	–	13	9	100	0	70
L_1	–	5.3	44	0	100	39	–	6.2	35	0	100	46
L_∞	–	3.2	26	–	–	46	–	5.1	32	–	–	56
Huber in C_γ	1.5	2.3	44	96	4	27	1.5	3.5	29	95	5	32
Huber in E_γ	0.18	0.08	92	43	57	9	0.07	0.16	87	27	73	9

Table A.2: Parameters of the CO ($\gamma^* = 0.05$) and PD ($\gamma^* = 0.003$) ARMA models and Huberian variance of each parameter λ^H.

			CO left		
i	1	2	3	4	5
a_i	$-0,877$	$-0,152$	$0,173$	$-0,215$	$0,073$
c_i	$-0,236$	$-0,065$	$0,141$	$-0,098$	-
$\lambda^H_{a_i}$	0.0021	0.0032	0.0015	0.0035	0.0026
$\lambda^H_{c_i}$	0.0012	0.0075	0.0056	0.0074	-
			PD left		
i	1	2	3	4	5
a_i	$-0,712$	$-0,022$	$-0,018$	$-0,181$	$-0,060$
c_i	$-0,166$	$0,119$	$0,160$	$0,133$	-
$\lambda^H_{a_i}$	0.0031	0.0022	0.0095	0.0015	0.0086
$\lambda^H_{c_i}$	0.0002	0.0005	0.0066	0.0024	-

Table A.3: Coefficients A_m^N in the covariance matrix of the CO ($\gamma^* = 0.05$) and PD ($\gamma^* = 0.003$) ARMA models.

CO left											
m	0	1	2	3	4	5	6	7	8	9	10
A_m^N	1	0.91	0.86	0.74	0.62	0.45	0.33	0.22	0.19	0.11	0.09
PD left											
m	0	1	2	3	4	5	6	7	8	9	10
A_m^N	1	0.94	0.81	0.71	0.63	0.51	0.41	0.29	0.18	0.10	0.08

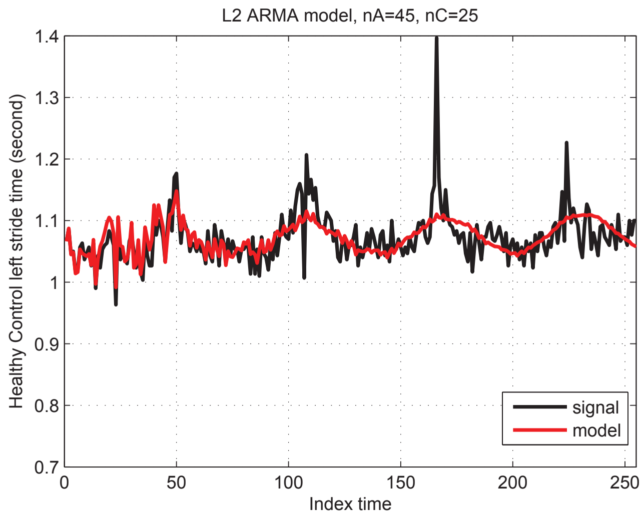

Figure A.3: Gaussian ARMA model of the left STS (red line) vs CO real signal (black line). $n_A = 45$, $n_C = 25$, $Fit = 9.5\%$.

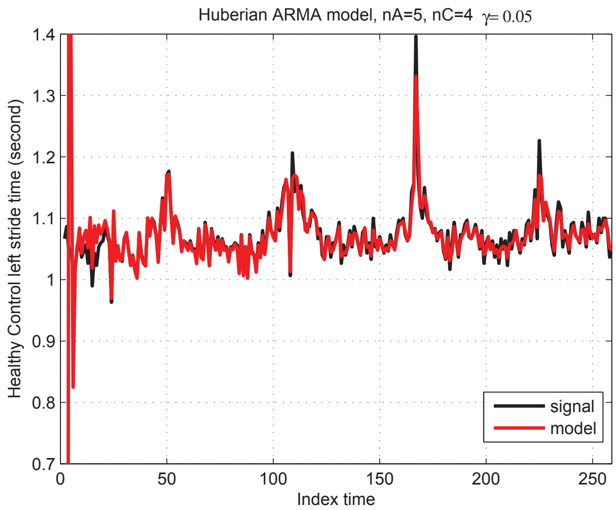

Figure A.4: Huberian ARMA model of the left STS (red line) vs CO real signal (black line). $n_A = 5$, $n_C = 4$, $Fit = 82.7\%$, $\gamma = 0.05$, $N = 253$.

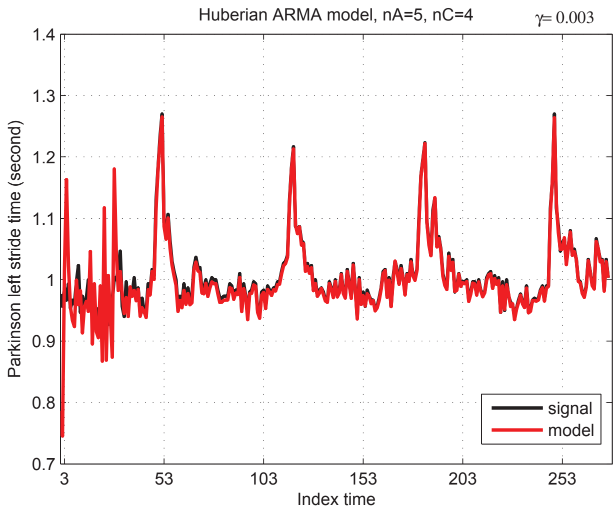

Figure A.5: Huberian ARMA model of the left STS (red line) vs PD real signal (black line). $n_A = 5$, $n_C = 4$, $Fit = 82.8\%$, $\gamma = 0.003$, $N = 288$.